BDT – cs523

Extra Credit Lab

- Submit your *own work* on time. No credit will be given if the lab is submitted after the due date.
- Note that the completed lab should be submitted in .zip format only.

Input dataset for this lab is same as Lab3.
Submit all the java files and output files.

1. **[6] Write a MapReduce java program to produce output as follows:**

011990-99999	**100**	**1950**
011990-99999	**80**	**1901**
011990-99999	**70**	**1930**
……	**…**	**…**
012650-99999	**120**	**1960**
012650-99999	**100**	**2015**
……		

Note that this output is same as input but only arranged in different format as shown above. (No need to find avg or max temperature here!)

In this output, the first column represents the stationID in ascending order, second column represents temperature in descending order and the last column is the year (no order defined for year).

2. **[4] This question requires some more research on your part.**
Modify the above program so that the output file name will be "StationTempRecord" instead of "part-r-00000".
Follow the following directions for this question.
    a. Do not "rename" the output folder to be "StationTempRecord".
    b. If your output folder name is "StationTempRecord-r-00000", then it's not an acceptable answer.
    c. Programmatically the output file name should be created as "StationTempRecord" and nothing else!

Remember to provide the command to run the above program in pseudo-distributed mode and paste screenshots wherever applicable.